

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 20, 2002, 14:47:28 ; Search time 355 Seconds

(without alignments)
3391.245 Million cell updates/sec

Title: US-09-802-285-2

Perfect score: 2043
Sequence: 1 MKKQILYLVQQLFLCSAY.....VGNSTVPTVYNSLSESTAR 384

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ .p2n.model -DEV=xih
-O=/cgn2_1/USPTO.spool/US09802285/runat.15112002.114655.18334/app-query.fasta-1.583
-DB=N_Geneseq.101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=trans40.cdi
-LIST=45 -DOCLIG=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEXT=pcr -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09802285 -GCGN.1.1.125 -runat.15112002.114655.18324 -ICPU=3
-NO_XLPY -NO_MMAP -LARGEOUTPUT -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

N_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|-------------|---------------------|
| 1 | 2043 | 100.0 | 1379 | 18 AAT87488 | Mutant polysacchar |
| 2 | 2043 | 100.0 | 1379 | 21 AA51243 | F. heparinum hepar |
| 3 | 2043 | 100.0 | 1379 | 23 AB150563 | Flavobacterium hep |
| 4 | 2040 | 99.9 | 1379 | 14 AAQ1046 | Heparinase. Flavo |
| 5 | 116 | 5.7 | 2538 | 24 ABN70382 | Streptococcus poly |
| 6 | 116 | 5.7 | 2562 | 24 ABN67361 | Streptococcus poly |
| 7 | 111 | 5.4 | 2365589 | 24 ABA90521 | Genomic sequence o |
| 8 | 107 | 5.2 | 2178 | 20 AA212261 | Neisseria gonorrhoe |
| 9 | 105 | 5.1 | 2178 | 20 AA212259 | Neisseria meningit |
| 10 | 105 | 5.1 | 2178 | 20 AA212260 | Neisseria meningit |
| 11 | 105 | 5.1 | 47475 | 21 AA481465 | Neisseria meningit |
| 12 | 105 | 5.1 | 349980 | 21 AAF21612 | Neisseria meningit |
| 13 | 105 | 5.1 | 837096 | 21 AA481489 | N. meningitidis pa |
| 14 | 103 | 5.0 | 5069 | 22 AA452292 | Human polynucleoti |
| 15 | 103 | 5.0 | 5151 | 23 ABV22238 | Human prostate exp |
| 16 | 103 | 5.0 | 5151 | 23 ABV28076 | Human prostate exp |
| 17 | 103 | 5.0 | 5465 | 22 AAS60778 | Human cancer agent |
| 18 | 102.5 | 5.0 | 6300 | 21 AA451864 | B. subtilis opp op |
| 19 | 102 | 5.0 | 1719 | 24 AB068105 | Listeria monocytog |
| 20 | 102 | 5.0 | 1146 | 22 AB069946 | Listeria monocytog |
| 21 | 101.5 | 5.0 | 1146 | 22 AAH01425 | Klebsiella pneumon |
| 22 | 101.5 | 5.0 | 6135 | 24 AB069931 | Listeria monocytog |
| 23 | 101 | 4.9 | 1707 | 20 AA220032 | Enterococcus faeca |
| 24 | 101 | 4.9 | 1707 | 20 ABN98017 | E faecalis EF018 g |
| 25 | 101 | 4.9 | 10955 | 20 AA131173 | Enterococcus faeca |
| 26 | 98 | 4.8 | 1653 | 24 ABN67396 | Streptococcus poly |
| 27 | 97.5 | 4.8 | 3227 | 24 ABK74276 | Bacillus lichenillo |
| 28 | 97 | 4.7 | 3721 | 20 AA13348 | Enterococcus faeca |
| 29 | 96 | 4.7 | 2070 | 23 AB110733 | Drosophila melanog |
| 30 | 95 | 4.7 | 1461 | 22 AA44660 | Equine infectious |
| 31 | 95 | 4.7 | 4768 | 22 AA479604 | Virulence gene #24 |
| 32 | 95 | 4.7 | 8255 | 22 AAH44659 | Equine infectious |
| 33 | 95 | 4.7 | 213251 | 22 AB067193 | Listeria innocua c |
| 34 | 94.5 | 4.6 | 1540 | 20 AA220033 | Enterococcus faeca |
| 35 | 94.5 | 4.6 | 1340 | 24 ABN98018 | E faecalis EF018 g |
| 36 | 94.5 | 4.6 | 29069 | 21 AA481497 | N. meningitidis pa |
| 37 | 94.5 | 4.6 | 349980 | 21 AAF21609 | Neisseria meningit |
| 38 | 94.5 | 4.6 | 1437668 | 21 AA481490 | N. meningitidis B |
| 39 | 94 | 4.6 | 1664976 | 19 AAV21209 | Methanococcus jam |
| 40 | 93.5 | 4.6 | 4557 | 24 ABK75361 | Bacillus lichenillo |
| 41 | 93 | 4.6 | 1149 | 21 AA44838 | Arabidopsis thalia |
| 42 | 93 | 4.6 | 2313 | 24 AB069770 | Listeria innocua D |
| 43 | 93 | 4.6 | 3108 | 21 AA479734 | Eucalyptus grandis |
| 44 | 93 | 4.6 | 684767 | 24 AB067196 | Listeria innocua c |
| 45 | 92.5 | 4.5 | 1580 | 19 AAV19363 | Cyclomaltodextrin |

ALIGNMENTS

RESULT 1
AAT87488
ID AAT87488 standard; DNA: 1379 BP.

AC AAT87488;
DT 02-FEB-1998 (first entry)

DE Mutant polysaccharide lyase derived from F. heparinum heparinase.

XX Polysaccharide: lyase; Flavobacterium heparinum; heparinase; mutant;

KW heparine; ds.
XX Flavobacterium heparinum.

OS
XX
FH Key Location/Qualifiers
FT 173..1327
FT /*tag= a

FT mat-peptide 236..1324
 FT /tag- b
 XX MO9716556-A1.
 XX PD 09-MAY-1997.
 XX PF 30-OCT-1996; 96MO-US17310.
 XX PR 30-OCT-1995; 95US-0008069.
 XX (MAST) MASSACHUSETTS INST TECHNOLOGY.
 XX Cooney CL, Ernst S, Godavarti R, Langer R, Sasisekharan R;
 PI Venkataraman G;
 DR WPI, 1997-272124/24.
 DR P-PSDB; AAMW28545.
 XX New polysaccharide lyase derived from Flavobacterium heparinum
 PT heparinase - by specified mutation(s), used to remove heparin from
 PT blood and to digest heparin to low molecular weight fragments
 PS Claim 15; Page 53-57; 73pp; English.
 CC The new polysaccharide lyase has the amino acid sequence encoded
 CC by the sequence given in AAT87488 with at least one of the
 CC following amino acid substitutions:
 CC (a) Cys135 replaced by Asp, Glu, Ser, Thr or His,
 CC (b) a conservative substitution in the Cardin-Weintraub-like
 CC heparin-binding sequence (HBS) XBXXHXB (positions 197-205
 CC or 208-212; B- any basic residue, X- any hydrophobic or
 CC other residue) with a residue that conforms to the HBS,
 CC (c) a conservative substitution in a EF-hand-like calcium
 CC binding site (CBS), position 206-220, with a residue that
 CC conforms to CBS,
 CC (d) a (non-)conservative substitution of a P81, P82 or P83 beta-
 CC sheet domain, provided this preserves the parallel beta-
 CC helix structure,
 CC (e) non-conservative substitution of Cys297,
 CC (f) deletion of one or more N- or C-terminal residues, provided
 CC the parallel beta-helix structure is retained,
 CC (g) substitution of His203 by Asp, Glu, Ser, Thr or Cys,
 CC (h) substitution of Lys198, 199, 205, 208, 211 or 214 with
 CC a small (non-)polar or acidic amino acid
 CC (i) substitution of Phe197, Asn200, Asp204, Glu207, Asp210,
 CC Asp212 or Gly213 by a small (non-)polar amino acid,
 CC (j) non-conservative substitution of Ser39.
 CC The mutant lyase is used to remove heparin from fluids, esp.
 CC circulating blood. The lyase may have (i) altered enzymatic
 CC activity compared with the native enzyme, or (ii) same activity
 CC but reduced immunogenicity (particularly where Ser39 is replaced).
 CC Also they may be made more stable and easier to produce and purify
 CC (pref. by removal of crosslinkable Cys residues).
 XX SQ Sequence 1379 BP; 465 A; 308 C; 281 G; 325 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 6,4e-213 Length: 1379
 Score: 2043.00 Matches: 384
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 18 Gaps: 0
 US-09-802-285-2 (1-384) x AAT87488 (1-1379)
 QY 1 MetLysGlnIleuLeuTyrLeuIleValLeuGlnIleuPheLeuCysSerAlaTyr 20
 DB 173 ATGAAAAAACAATTCTATCTCATCTGACTTACGACAACTGCTTCTGTCGGCTTAC 232
 QY 21 AlAGlnIleuLysSerGlyAsnIleProTyrArgValAsnValGlnAlaAspSerAla 40

DB 233 GCCCAGCAAAAAATCCCGTAACATCCCTTACCGGGTAATGTGACGGCCGACACTGCT 292
 QY 41 LysGlnLysAlaIleIleAspAsnLysTyrValAlaValGlyIleAsnLysProTyrAla 60
 DB 293 AAGCAGAAAGCGCATTTATGACAAACAATGGGTGGAGTAGGACATCAATTAACCTTATGCA 352
 QY 61 LeuGlnTyrAspAspLysLeuArgPheAsnGlyLysProSerTyrArgPheGluLeuLys 80
 DB 353 TTACAATATGACATTAACGTGGCTTAAATGGAAAAACATCCATCCATGCTTGGCTTAA 412
 QY 81 AlaGlnAspAsnSerLeuGlnGlyTyrAlaAlaGlyIleGlyTyrGlnTyrGluLeu 100
 DB 413 GCCGAGACAAATTCGCTTGAAGTTATGCTGCGAGGAAACAAAGGCCGCTACAGAAATG 472
 QY 101 SerTyrSerTyrAlaThrThrAsnAspPheLysLysPheProSerValTyrGlnAsn 120
 DB 473 TCGTACAGCATATGCAACCAATGATTTTAAAGAAATTTCCCCAGCGATACCAAAAT 532
 QY 121 AlaGlnLysLeuLysThrValTyrHisTyrGlyLysGlyIleGlyGlnGlnGlySerSer 140
 DB 533 GCGCAAAAGCTAAAACCGTTTATCATACGGCAAAAGGATTTGTAAACAGGGGACCTCC 592
 QY 141 ArgSerTyrThrPheSerValTyrIleProSerSerPheProAspAsnAlaThrThrIle 160
 DB 593 CGCAGCATATACCTTTTCAGGTACATACCTCCCTCCCGACAAATGCACTACTATT 652
 QY 161 PheAlaGlnTyrPheIleGlyAlaProSerArgThrLeuValAlaThrProGluGlyIle 180
 DB 653 TTTGGCCCAATGGCATGTGTCACACAGACAGAACCTTGTAACACAGAGGAGAAAT 712
 QY 181 LysThrLeuSerIleGluGlnPheLeuAlaLeuTyrAspArgMetIlePheLysAsn 200
 DB 713 AAACACTGAGCATTAAGAGGTTTGGCTTATACGACCGCATGATCTTCAAAAAAAT 772
 QY 201 IleAlaHisAspLysValGluLysLysAspGlyLysIleThrTyrValAlaGly 220
 DB 773 ATGCCCATATTAAGTTGAAAAAAGAAAGAGAGGAAATTTACTATGACCGCA 832
 QY 221 LysProAsnGlyTyrPlyValGluGlnGlyGlyTyrProThrLeuAlaPheGlyPheSer 240
 DB 833 AACCCAAATGGCTGGAAGGAGAGCAAGAGGTATCCACGCTGCGCTTGGTTTTC 892
 QY 241 LysGlyTyrPheTyrIleLysAlaAsnSerAspArgGlnTyrPheThrAspLysAlaAsp 260
 DB 893 AAAGGATTTTATCAATCAAGCAAACTCCGACCGCATGGCTTACCGCAAAAGCGAC 952
 QY 261 ArgAsnAsnAlaAsnProGluAsnSerGluValIleLysProTyrSerSerGlyTyrLys 280
 DB 953 CGTAACATGCCAATCCCGAGATATGAAATGAAAGCCCTATCTCGGAATACAA 1012
 QY 281 ThrSerThrIleAlaTyrLysMetProPheAlaGlnPheProLysAspCysTyrIleThr 300
 DB 1013 ACTTCAACCATTCCTATATAAATGGCCCTTGGCCAGTCTTCAAGATTCGTGATTA 1072
 QY 301 PheAspValAlaIleAspTyrThrLysTyrGlyGlnAlaAsnThrIleLeuLysPro 320
 DB 1073 TTGATGTCCGCATAGACTGAGCAATATGAAAGAGGCCAATTAATTTGAAACCC 1132
 QY 321 GlyLysLeuAspValMetMetThrTyrThrLysAsnLysLysProGlnLysAlaHisIle 340
 DB 1133 GGTAAAGTGTATGTATGATGACTTATACCAAAATTAAGAAACACAAAGCGCATATC 1192
 QY 341 ValAsnGlnGlnIleuLysIleGlyArgAsnAspAspLysTyrTyrPheLysPhe 360
 DB 1193 GTAAACAGCAGCAAAATCCGATCGGACGTAAACGATGACATGCTTACTCAATTT 1252
 QY 361 GlyIleTyrArgValGlyAsnSerThrValProValThrTyrAsnLeuSerGlyTyrSer 380
 DB 1253 GGAATTTACGGGTGTGTAACAGACAGGTCCGGTTACTTATTAACGTGAGCGGTACAC 1312
 QY 381 GluThrAlaArg 384
 DB 1313 GAAACTGCCAGA 1324

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame-plus_p2n model

Run on: November 20, 2002, 14:54:03 ; Search time 52 Seconds
(Without alignments)
2264.690 Million cell updates/sec

Title: US-09-802-285-2
Perfect score: 2043
Sequence: 1 MKKQILYLVQLQFLCSAY.....VGNSTVPTVYNSGSETAR 384

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 88:724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_{p2n}.model -DEV=xlh
-O=/cgn2_1/USPTO/US09802285/runat_15112002_114657_18362/app_query.fasta_1.583
-DB=Issued_Patents_NA -QEXT=fastap -SUFFIX=ini -MINMATCH=0.1 -IOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09802285.ecgn1.1.17.ernat.15112002.114657.18362 -MCPU=6 -ICPU=3
-NO_XIPX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database:

- Issued Patents_NA:*
- 1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|------------------------|
| 1 | 2043 | 100.0 | 1379 | 1 | US-07-783-706-1 / 7020 |
| 2 | 2043 | 100.0 | 1379 | 5 | PCT-US92-09124-1 |
| 3 | 2043 | 100.0 | 1407 | 4 | US-09-066-481-1 |
| 4 | 2038 | 99.8 | 1379 | 2 | US-08-445-342A-1 |
| 5 | 101 | 4.9 | 1707 | 4 | US-09-071-035-65 |
| 6 | 94.5 | 4.6 | 1540 | 4 | US-09-071-035-67 |
| 7 | 92.5 | 4.5 | 3695 | 4 | US-09-071-035-419 |
| 8 | 92.5 | 4.5 | 3840 | 4 | US-09-071-035-417 |
| 9 | 92 | 4.5 | 11864 | 4 | US-08-961-527-61 |
| 10 | 91 | 4.5 | 4661 | 4 | US-09-221-017B-970 |
| 11 | 90.5 | 4.4 | 2630 | 3 | US-08-668-286-6 |
| 12 | 90.5 | 4.4 | 2630 | 4 | US-09-469-253-6 |

| | | | | | | |
|----|------|-----|------|---|---------------------|--------------------|
| 13 | 90.5 | 4.4 | 2630 | 4 | US-09-642-146-6 | Sequence 6, Appli |
| 14 | 90 | 4.4 | 980 | 4 | US-09-221-017B-589 | Sequence 589, App |
| 15 | 90 | 4.4 | 2860 | 4 | US-09-221-017B-871 | Sequence 871, App |
| 16 | 90 | 4.4 | 6885 | 1 | US-08-370-311-9 | Sequence 9, Appli |
| 17 | 90 | 4.4 | 6895 | 2 | US-08-353-485-9 | Sequence 9, Appli |
| 18 | 89.5 | 4.4 | 3827 | 2 | US-08-447-031A-1 | Sequence 1, Appli |
| 19 | 89.5 | 4.4 | 4612 | 2 | US-08-447-031A-8 | Sequence 8, Appli |
| 20 | 89 | 4.4 | 2191 | 4 | US-09-071-035-363 | Sequence 363, App |
| 21 | 89 | 4.4 | 2334 | 4 | US-09-071-035-361 | Sequence 361, App |
| 22 | 89 | 4.4 | 4310 | 3 | US-09-008-172-1 | Sequence 1, Appli |
| 23 | 89 | 4.4 | 4310 | 4 | US-09-210-361-5 | Sequence 5, Appli |
| 24 | 89 | 4.4 | 4310 | 4 | US-09-740-274-5 | Sequence 5, Appli |
| 25 | 88 | 4.3 | 4460 | 3 | US-09-007-999-1 | Sequence 1, Appli |
| 26 | 88 | 4.3 | 4460 | 4 | US-09-210-361-1 | Sequence 1, Appli |
| 27 | 88 | 4.3 | 4460 | 4 | US-09-740-274-1 | Sequence 7, Appli |
| 28 | 87.5 | 4.3 | 1169 | 4 | US-09-386-642-7 | Sequence 7, Appli |
| 29 | 87.5 | 4.3 | 2061 | 1 | US-08-204-656B-9 | Sequence 9, Appli |
| 30 | 87.5 | 4.3 | 2061 | 1 | US-08-470-702-5 | Sequence 5, Appli |
| 31 | 87.5 | 4.3 | 2061 | 1 | US-08-467-831-5 | Sequence 5, Appli |
| 32 | 87.5 | 4.3 | 4344 | 2 | US-08-532-547-4 | Sequence 4, Appli |
| 33 | 87.5 | 4.3 | 4344 | 2 | US-08-379-656B-4 | Sequence 4, Appli |
| 34 | 87.5 | 4.3 | 4344 | 3 | US-08-455-838-4 | Sequence 4, Appli |
| 35 | 87.5 | 4.3 | 4344 | 3 | US-09-019-809-4 | Sequence 4, Appli |
| 36 | 87.5 | 4.3 | 4344 | 4 | US-09-471-177-4 | Sequence 4, Appli |
| 37 | 87.5 | 4.3 | 9711 | 4 | US-08-451-715A-7 | Sequence 167, App |
| 38 | 87 | 4.3 | 2973 | 1 | US-08-476-062A-40 | Sequence 4, Appli |
| 39 | 87 | 4.3 | 3533 | 2 | PCT-US96-01314-40 | Sequence 40, Appli |
| 40 | 87 | 4.3 | 3533 | 6 | 5424399-1 | Sequence 40, Appli |
| 41 | 87 | 4.3 | 6836 | 4 | US-08-976-259-73 | Sequence 73, Appli |
| 42 | 86.5 | 4.2 | 1585 | 4 | US-09-071-035-43 | Sequence 43, Appli |
| 43 | 86.5 | 4.2 | 1659 | 4 | US-09-071-035-41 | Sequence 41, Appli |
| 44 | 86.5 | 4.2 | 1983 | 4 | US-09-134-001C-1524 | Sequence 1524, Ap |
| 45 | 86.5 | 4.2 | 1983 | 4 | | |

ALIGNMENTS

PD Feb 3 '98

RESULT 1
US-07-783-706-1
Sequence 1, Application US/07783706
Patent No. 5714376
GENERAL INFORMATION:
APPLICANT: Sasisekharan, Ramnath
APPLICANT: Moreman, Kelley L.
APPLICANT: Cooney, Charles L.
TITLE OF INVENTION: The Heparinase Gene from Flavobacterium
TITLE OF INVENTION: Heparin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street, Suite 3100
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/783,706
FILING DATE: 19911023
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MITS546
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 1:

| | | | |
|----|------|---|------|
| Db | 893 | AAAGGATTTTACATCAGCAAGCAAACTCGACGCGGACATGGCTTACCGACAAAGCCAC | 952 |
| Qy | 261 | ArgsnasnaalaaasnProgluaasSerGluaValMetLysProTyrSerSerLutIrrLys | 280 |
| Db | 953 | CGTAAACATCCCAATCCCGGAATAGTGAAGTAATCAACCCCTATTCTCGGAATACAA | 1012 |
| Qy | 281 | ThSerThrIleAlaTyrLysMetProPheAlaGlnPheProLysAspCysTrpIleThr | 300 |
| Db | 1013 | ACTCAACCACTGGCTATATAAATGGCTTTGGCCAGTTCCCTAAAGATTGCGGATTACT | 1072 |
| Qy | 301 | PheaspAlaIleAlaSprprhrLysTyrGlyLysGluAlaAsnThrIleLeuLysPro | 320 |
| Db | 1073 | TTTTATGTCCGCATAGCTGGACGAAATATGGAAAAAGGCCAAATACATTTTGAACCC | 1132 |
| Qy | 321 | GlyLysLeuAspValMetMetThrTyrThrLysAsnLysLysProGluLysAlaHisIle | 340 |
| Db | 1133 | GGTAGCTGGATGTGATGATGACTTATACCAAGAAATAAAGAACCCCAAAAAAGCGATATC | 1192 |
| Qy | 341 | ValasnsInGlnIuIleLeuIleGlyLysGAsnAspAspAspGlyTyrTyrPheLysPhe | 360 |
| Db | 1193 | GTAACCAAGAGGAATTCCTGATCGGACCTATACGATGACGATGACGCTATTACTTCAAAATTT | 1252 |

381 GlutThraLarg 384
 1313 GAAACTGCCAGA 1324
 Db

RESULT 2
PCT-US92-09124-1

; Sequence 1, Application PC/TUS9209124

GENERAL INFORMATION:

APPLICANT: Massachusetts Institute, of Technology

TITLE OF INVENTION: The Heparinase Gene from *Flavobacterium*

TITLE OF INVENTION: Hep

NUMBER OF SEQUENCES: 2

ADDRESS: K17Dattick

STREET: 1100 Peachtree

CITY: Atlanta

STATE: Georgia

; COUNTRY: / US

ZIP: 30309-4530 ;

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-

SOFTWARE: Patent In Re

CURRENT APPLICATION DATA ;

APPLICATION NUMBER: F

FILING DATE: 19921022

CLASSIFICATION:

NAME: Pabst Patricia

REGISTRATION NUMBER:

REFERENCE/DOCKET NUMBER

TELECOMMUNICATION INFORMATION ;

TELEPHONE: 404-815-63

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS
LENGTH: 1379 bases

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANAL-SENSE: NO
ORIGINAL SOURCE:

ORIGINAL SOURCE: ORGANISM: *Flavobacter*

PCT-US92-09124-1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

OM protein - protein search, using sw model

Run on: November 20, 2002, 14:41:43 : Search time 45 Seconds
(without alignments) 858.503 Million cell updates/sec

Title: US-09-802-285-2

Sequence: 1 MKKQILYLVVQQLFLCSAY.....VGNSTVPVTYNI.SGSETAR 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 2043 | 100.0 | 384 | 2 A47479 | heparin lyase (EC |
| 2 | 111 | 5.4 | 398 | 2 C66650 | hypothetical prote |
| 3 | 107 | 5.2 | 409 | 2 E69847 | hypothetical prote |
| 4 | 105 | 5.1 | 363 | 2 AE0616 | cuter membrane pro |
| 5 | 105 | 5.1 | 725 | 2 D81976 | probable ferric si |
| 6 | 105 | 5.1 | 725 | 2 H81030 | ionb-dependent rec |
| 7 | 103 | 5.0 | 1452 | 1 S17669 | protein-tyrosine-p |
| 8 | 102.5 | 5.0 | 545 | 2 A38447 | clipeptide ABC t |
| 9 | 102.5 | 5.0 | 1243 | 2 S60138 | sex factor aggrega |
| 10 | 102 | 5.0 | 572 | 2 AD1209 | N-acetylmuramyl-L |
| 11 | 101.5 | 5.0 | 2044 | 2 AB1180 | probable peptidogl |
| 12 | 100.5 | 4.9 | 381 | 2 S45109 | beta-lactamase (EC |
| 13 | 99.5 | 4.9 | 612 | 2 G83307 | hypothetical prote |
| 14 | 99 | 4.8 | 363 | 2 S43159 | outer membrane por |
| 15 | 98.5 | 4.8 | 416 | 2 A97093 | probable membrane |
| 16 | 98.5 | 4.8 | 2523 | 2 T18477 | hypothetical prote |
| 17 | 98 | 4.8 | 515 | 2 T09272 | probable membrane |
| 18 | 98 | 4.8 | 546 | 2 D66811 | hypothetical prote |
| 19 | 98 | 4.8 | 1452 | 1 S17670 | probable tail comp |
| 20 | 97.5 | 4.8 | 537 | 2 A35400 | mannosyl-9-glycopr |
| 21 | 97.5 | 4.8 | 1873 | 2 T30944 | protein-tyrosine-p |
| 22 | 97 | 4.7 | 790 | 2 T49542 | surface protein T6 |
| 23 | 97 | 4.7 | 1628 | 2 E90538 | xylan 1,4-beta-xy |
| 24 | 96.5 | 4.7 | 1009 | 2 C64483 | hypothetical prote |
| 25 | 95.5 | 4.7 | 381 | 2 S38196 | beta-lactamase (EC |
| 26 | 95.5 | 4.7 | 953 | 2 S55156 | probable membrane |
| 27 | 95 | 4.7 | 684 | 2 S41788 | endo-1,4-beta-xyla |
| 28 | 95 | 4.7 | 1077 | 2 D71660 | hypothetical prote |
| 29 | 95 | 4.7 | 1092 | 2 S42798 | fibronectin-bindin |

| | | | | | |
|----|------|-----|------|----------|---------------------|
| 30 | 94.5 | 4.6 | 381 | 2 S08296 | beta-lactamase (EC |
| 31 | 94.5 | 4.6 | 421 | 2 G97319 | hypothetical prote |
| 32 | 94.5 | 4.6 | 657 | 2 E81119 | tail fibre protein |
| 33 | 94.5 | 4.6 | 718 | 1 ALBSMX | cyclomalodextrin |
| 34 | 94 | 4.6 | 761 | 1 B64506 | DNA topoisomerase |
| 35 | 94 | 4.6 | 823 | 2 S14055 | nucleoskeletal-lik |
| 36 | 93.5 | 4.6 | 610 | 2 T22909 | hypothetical prote |
| 37 | 93.5 | 4.6 | 710 | 2 S22673 | ferritinamine recep |
| 38 | 93 | 4.6 | 770 | 2 A11769 | autolysin, amidase |
| 39 | 93 | 4.6 | 1750 | 2 H64432 | ribonucleoside-tri |
| 40 | 92.5 | 4.5 | 315 | 2 AC0392 | probable exported |
| 41 | 92.5 | 4.5 | 422 | 2 F70018 | multiple sugar-din |
| 42 | 92.5 | 4.5 | 1520 | 2 T44231 | hypothetical prote |
| 43 | 92 | 4.5 | 500 | 2 B43577 | perfringolysin O p |
| 44 | 92 | 4.5 | 515 | 2 T09203 | probable tail comp |
| 45 | 92 | 4.5 | 652 | 2 G95177 | hypothetical prote |

ALIGNMENTS

RESULT 1

A47479 heparin lyase (EC 4.2.2.7) I precursor - Flavobacterium heparinum

C:Species: Flavobacterium heparinum

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999

C:Accession: A47479; S65553

R:Sasisekharan, R.; Bulmer, M.; Moremen, K.W.; Cooney, C.L.; Langer, R.

Proc. Natl. Acad. Sci. U.S.A. 90, 3660-3664, 1993

A:Title: Cloning and expression of heparinase I gene from Flavobacterium heparinum.

A:Reference number: A47479; PMID:93234557; PMID:8475114

A:Accession: A47479

A:Status: preliminary

A:Molecule type: DNA; protein

A:Residues: 1-384 <SAS>

A>Note: sequence extracted from NCBI database (NCBI:129766, NCBI:129769)

R:Ernst, S.; Venkataraman, G.; Winkler, S.; Godavarti, R.; Langer, R.; Cooney, C.L.;

Biochem. J. 315, 589-597, 1996

A:Title: Expression in Escherichia coli, purification and characterization of heparin

A:Reference number: S65553; PMID:96202015; PMID:8615834

A:Accession: S65553

A:Molecule type: protein

A:Residues: 22-27 <ERN>

C:Superfamily: Flavobacterium heparinum heparin lyase I

C:Keywords: carbon-oxygen lyase

Query Match Best Local Similarity 100.0%; Score 2043; DB 2; Length 384;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|--|
| QY | 1 | MKKQILYLVVQQLFLCSAYAAQOKSGNIPYRVNQADSAKOKAIIDNKVAVGINKPYA 60 | |
| DB | 1 | MKKQILYLVVQQLFLCSAYAAQOKSGNIPYRVNQADSAKOKAIIDNKVAVGINKPYA 60 | |
| QY | 61 | LOYDCKLRNKGPSYRFELKADNSLEGYAAGTYGRFELSYVATNDKFFPSVYON 120 | |
| DB | 61 | LOYDCKLRNKGPSYRFELKADNSLEGYAAGTYGRFELSYVATNDKFFPSVYON 120 | |
| QY | 121 | AKKLTIVYHGGIGEOGSSRSYTPSPSPNATTFQWVGAPRTLVATPEGFI 180 | |
| DB | 121 | AKKLTIVYHGGIGEOGSSRSYTPSPSPNATTFQWVGAPRTLVATPEGFI 180 | |
| QY | 181 | KTLSTEEFLATYDRIAFKKNIAHDKVEKKDKGKITYVAGKNGKVEGQGYTLAFGS 240 | |
| DB | 181 | KTLSTEEFLATYDRIAFKKNIAHDKVEKKDKGKITYVAGKNGKVEGQGYTLAFGS 240 | |
| QY | 241 | KGYFIKANSDDQWLTDAKDRNNANPENSEVAKPYSEYKSTIAYKMPFAOPKDCWIT 300 | |
| DB | 241 | KGYFIKANSDDQWLTDAKDRNNANPENSEVAKPYSEYKSTIAYKMPFAOPKDCWIT 300 | |
| QY | 301 | FDVAIDWTKYGEANTILKPGKLDVAMTYTKKKKQKAHIVNQOILLGRNDDGYERF 360 | |
| DB | 301 | FDVAIDWTKYGEANTILKPGKLDVAMTYTKKKKQKAHIVNQOILLGRNDDGYERF 360 | |

QY 361 GIVRGVSTVPYVYNYLSGSETAR 384
 DB 361 GIVRGVSTVPYVYNYLSGSETAR 384

RESULT 2

hypothetical protein rygd (imported) - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 24-Aug-2001
 C:Accession: C86650

R:Bolotin, A.; Wincker, P.; Manger, S.; Jallion, O.; Mairme, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
 A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: C86650
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-398 <STO>
 A:Cross-references: GB:AEO05176; PID:g12723058; PIDN:AAK04301.1; GSPDB:GN00146
 A:Experimental source: strain IL1403

C:Genetics:
 A:Gene: rygd
 C:Superfamily: Streptococcus mutans ABC transporter rygDc: ATP-binding cassette homology

Query Match 5.4%; Score 111; DB 2; Length 398;
 Best Local Similarity 22.9%; Pred. No. 0.55;
 Matches 82; Conservative 43; Mismatches 125; Indels 108; Gaps 20;

QY 41 KQAIIDNKKVVA-----VGINKPYALOYDOKLRNGKPSYRFELEKADNSIEGAAETK 95
 DB 84 KGVAVYNGKLVSTIEAGVGNP-----ELTGKENV--YLNQAML 120
 QY 96 G--RTELSSVATNTDFKPPSYQNAOKLKYHYHGKICEGSSRSYTFVSPSSF 153
 DB 121 GFSREISDMYDEIVDAEL--EEFNM-QRLK--NYSSGM-----QVRLAFSAIKNA-- 161
 QY 154 PDNATITFAOMHGAPSTVLATPEGEIKTISIEEFLLAYRMIFFKKNIANDYKCKDKG 213
 DB 168 -----RGDY--LVLEVLAVGDESPQRK--CNDYFPRKCKSG 200
 QY 214 KITTY-----AGK-----PNGMKVGGGPTLAFGSKGYF--IKAND----- 251
 DB 201 KITTLVTHMGAAKRYCNKAILLENGLIKSGVNDVANDNISLDNLC EYVARSDELIDV 260
 QY 252 -RQWLTDKADRNANPE-----NSEVMKPYSEVYKSTIAY-KMPEAOPKDCWITFDVAI 305
 DB 261 KEEMI-ENLEVNFINOOKIRPNDEVKREVSVDYKKDTEYTAPELTVDNRINWYNSLS 319
 QY 306 DMTRYGKEANTILPKGLD-----VMATYTKKKPKKAHYNOEILGRND--DD 354
 DB 320 DYLTGKGHNRATYCTKLDQVNNLKLQVSRKRDMEIAFADKYLINRNDISD 377

RESULT 3

hypothetical protein yjcm - Bacillus subtilis

C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 28-Jul-2000
 C:Accession: E69847

R:Kunert, F.; Ogasawara, N.; Muszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berti
 C: Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cinc
 A: Ehrlich, S.D.; Emerson, P.T.; Ertlin, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Flytz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, M.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E69847
 A:Status: preliminary
 A:Molecule type: nucleic acid sequence not shown; translation not shown

A:Residues: 1-409 <KUN>
 A:Cross-references: GB:Z99110; GB:AL009126; NID:92633472; PIDN:CAD13048.1; PID:e1183;
 C:Experimental source: strain 168

C:Genetics:
 A:Gene: yjcm
 C:Superfamily: Bacillus subtilis hypothetical protein yjcm

Query Match 5.2%; Score 107; DB 2; Length 409;
 Best Local Similarity 20.1%; Pred. No. 1.2;
 Matches 85; Conservative 56; Mismatches 155; Indels 126; Gaps 15;

QY 1 MKKQILYIVL-----QQLPLCSAVAQKSGNIPYRVNVAQDSAKQKAIIDNKW 50
 DB 1 MKKQLASIVLCISPLVSTNEVFATTKETKENVSTNEIGYE---KVKDMEQOK 56
 QY 51 VAVGINKPYALOYDOKLRNGKPSYRFELEKADNSIEGAAETKRTLSYATNTDF 110
 DB 57 YDFEHHTPLAHPRKSKL-----TDSQGV-----LLYSALSKVDK 91
 QY 111 KRPFSYQNAOKLKYHYHGKICEGSSRSYTFVSP--YIPSSFPDN----- 156
 DB 92 QKHFEFTY-DIDVKTSEKNNKLIYEAITRNVPFGEVKTYSLGNIKLEIKRTTSNKO 150
 QY 157 -----ATTFAOMHGAPSTVLATPEGEIKTIS----- 185
 DB 151 SSOSTNTSNLSFSQVKNSSNQLNVIPLSNSSQOLNTQLNNSSTLYQENDEEYSSDVKL 210
 QY 186 EEFLLAYDMIEFKKNIAHDKVEKDDKITYTAAGRNPKGVQGGPTLAFPSKGY-- 243
 DB 211 DEFLKTKQDEV--KNDKREKSEDKPVSNFNAANSNOLTV---ATSVKYS-GYAA 263
 QY 244 -----FYKANSRQWLTDKADRNANPENSEVMKPYSEVYKSTIAYKMPEAOPKDC 297
 DB 264 MKYAEYALKPKNNNYKYNDKQDNTNVSQALRGMRPFYEMKPYDA----- 311
 QY 298 WITFDVAIDMTYKKEANTILPKGL--DVMATYTKKKPKKAHYNOEILGRND 354
 DB 312 -----WVNAAGAFRSYILKAGIKMKTVSDTYSNVKLGIDYHYDHNK--IGLRAD 360
 QY 355 GX 356
 DB 361 GW 362

RESULT 4

outer membrane protein F precursor STY1002 (imported) - Salmonella enterica subsp. e

C:Species: Salmonella enterica subsp. enterica serovar typh
 A:Note: this species has also been called Salmonella typh
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AE0616

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 S.; Mout, S.; O'Gaora, P.
 Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica s
 A:Reference number: AB0502; PMID:11677608

A:Accession: AE0616
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-363 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD05399.1; PID:g16502161; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1002

C:Superfamily: outer membrane protein phoE

Thu